DOCUMENT RESUME

ED 145 698 FL 008 959

AUTHOR Johansson, Stig
TITLE Partial Dictation as a Test of Foreign Language Proficiency. Swedish-English Contrastive Studies, Report No. 3.
INSTITUTION Lund Univ. (Sweden). Dept. of English.
PUB DATE 73
NOTE 56p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Cloze Procedure; English (Second Language); Error Analysis (Language); Grammar; *Language Instruction; Language Laboratory Use; *Language Proficiency; Language Skills; *Language Tests; *Listening Comprehension; Listening Skills; *Listening Tests; Morphology (Languages); Phonology; Second Language Learning; Tape Recordings; Testing; Vocabulary
IDENTIFIERS *Dictation

ABSTRACT

This paper presents arguments in favor of using partial dictation as a test of foreign language proficiency. In this kind of test, subjects listen to recordings of material in the foreign language and are required to fill in missing words in a written version of the recordings. Partial dictation is preferable to ordinary dictation in that: (1) it makes possible the use of different voices and speech situations; (2) it is more economical; and (3) the testing situation is more natural, being subject to fewer interruptions. In addition, partial dictation appears to be reliable and to correspond very well to other measures of foreign language proficiency; it is easy to construct, administer, and score; and it is useful, not only as a measure of listening comprehension, but as a global estimate of language proficiency. An incomplete analysis is presented of errors found in connection with this test, on the levels of phonology, lexicology, and grammar. Appendices contain two sample passages used in a partial dictation test, and a listing of average scores for each item, from a random sampling of 20 subjects. (Author/AM)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

LUND UNIVERSITY

ED145698

Swedish-English Contrastive Studies

Partial Dictation as a Test of Foreign Language Proficiency

by

Stig Johansson

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Jan Svartik
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Report No 3

1973

Department of English
Lund University

Helgonabacken 14
S-223 62 Lund, Sweden

J. Svartvik gift
FL009759
Jan. 15, 1977

CONTENTS

		Page
1	AIM	1
2	BACKGROUND	2
	2.1 Dictation	2
	2.2 The cloze technique	5
3	PARTIAL DICTATION	9
	3.1 Reliability	9
	3.2 Validity	10
	3.3 Construction and administration	12
	3.4 Scoring system	14
4	ERROR ANALYSIS	18
	4.1 Types of errors in partial dictation	18
	4.1.1 Phoneme errors	18
	4.1.1.1 Omission	19
	4.1.1.2 Substitution	20
	4.1.1.3 Addition	24
	4.1.1.4 Reordering	24
	4.1.2 Lexical errors	25
	4.1.2.1 Omission	25
	4.1.2.2 Substitution	26
	4.1.2.3 Addition	28
	4.1.3 Grammatical errors	28
	4.1.3.1 Segmentation	28
	4.1.3.2 Reordering	30
	4.1.3.3 Restructuring	31
	4.1.3.4 Morphology	33
	4.1.3.5 Function words	35
	4.2 Levels of comprehension	36
	4.3 The linguistic evidence of comprehension errors	38
5	WHAT DOES PARTIAL DICTATION MEASURE?	40
	NOTES	43
	REFERENCES	47
	APPENDIX 1: The test	50
	APPENDIX 2: Discrimination of test items	53

1 AIM

In this paper arguments will be presented in favour of using partial dictation as a test of foreign language proficiency. In partial dictations test subjects listen to recordings of material in the foreign language and are required to fill in missing words in a written version of the recordings. It will be shown that such tests are reliable and correspond very well to other measures of foreign language proficiency and that they are, in addition, easy to construct, administer and score. The paper includes a classification of errors in a partial dictation.

2 BACKGROUND

Partial dictations are closely related to two other types of tests which have been used to measure foreign language proficiency: dictation and cloze tests. We shall discuss the use of partial dictation against the background of these testing techniques in order to (1) clarify the theoretical background of partial dictation, and (2) bring out the advantages of partial dictation over ordinary dictation and tests based on the cloze technique.

2.1 DICTATION

In ordinary dictation the teacher reads out passages in the foreign language and students are required to write them down. Valette (1967, p. 140) recommends the following technique in administering a dictation:

> First, the whole passage is read at normal speed. The students are told not to write, just to listen carefully. Then the passage is read a phrase at a time, with pauses during which the students write down what they have heard. At this time the teacher may read each phrase either once or twice, as long as he is consistent ... Finally, the entire passage is read again at normal speed, and the students are given a few minutes for final revision. It is imperative that the teacher never repeat a particular phrase at a student's request.

Valette (1967, p. 139) reports that dictation scores correlate very well with overall language achievement and refers to an earlier study on the subject (Valette 1964). In a recent article Oller (1971) has further demonstrated the remarkably high correlation between scores in dictations and other tests of foreign language proficiency.[1] After examining a test consisting of five parts (vocabulary, grammar,

composition, phonology and dictation), Oller concludes:

> It was discovered that the dictation correlated more highly with each other part of the test than did any other part. (p. 254)
>
> In short, the dictation clearly seems to be the best single measure of the totality of English-language skills being tested. (p. 255)

The correspondence between scores in dictations and other tests is not surprising, if we consider the process of speech perception. This has been convincingly argued by Oller (1971). In recent accounts of speech perception it is generally emphasized very strongly that the decoding of a spoken message is an active process in which the listener utilizes his knowledge of all aspects of the language system. Cf. the summary of the operations involved in speech decoding in Fry (1970, p. 51 f.):

1. Incoming acoustic information is processed by the acoustic decoding mechanism which works on a complex of cues appearing in the output.
2. Knowledge of the phonological constraints is combined with the results of acoustic processing; only a restricted set of phonemes is possible at any position.
3. The phoneme string is predicted on the basis of the items which are most probable acoustically and linguistically.
4. The phoneme string is segmented into morphemes by reference to knowledge of morpheme-forming rules.
5. Morpheme constraints are applied and morphemes identified by matching with the store of free and bound morphemes, beginning with the most probable item.
6. The morpheme string is formed into words matched with lexical items in the store, the choice being restricted by sequential dependencies.

7. The word sequence is formed which satisfies all syntactic and semantic constraints.

Fry points out that the "steps are inevitably set out one after the other but it is important to remember that they represent parallel processes with the very minimum of 'real time' lag between any level and the next immediately below" (p. 51). The fact that the listener utilizes his knowledge on all levels of the language is clearly borne out by an examination of errors in partial dictations (cf. Section 4).

Ordinary dictation is especially suitable as a test for beginning students. On more advanced levels the technique must be modified. Spolsky et al. (1968) suggested the use of different levels of white noise in a dictation test containing a series of isolated sentences. However, it has recently been shown (Johansson 1973) that performance in such a test may not be a true reflection of the ability to function in more realistic situations. Another way of modifying the testing technique is to concentrate the dictation to especially selected portions of the text, i.e. to use partial dictation. Partial dictation, as defined in this paper, also differs from ordinary dictation in the following respects:

a) The passages presented for dictation are played from a tape-recorder, not read out by the teacher.[2]
b) Students listen to the recordings only once.

Partial dictation is preferable to ordinary dictation in several respects. Since the passages to be written down are played from a tape-recorder, it is possible to use a variety of different voices and speech situations (recordings in different dialects and registers). Above all, partial dictation is more economical, since it can be administered more rapidly than ordinary dictation and can be concentrated on problem areas in the text. Finally, the

testing situation in partial dictation seems to be more natural, since the passages heard are not interrupted as often as in ordinary dictation.

2.2 THE CLOZE TECHNIQUE

The cloze technique was originally introduced as a means of measuring the readability of texts (Taylor 1953). In cloze tests every nth word in a sample of a text is deleted and replaced by a blank, and test subjects are asked to fill in the missing word. The score of a group of subjects is taken as a measure of the readability of the text. The cloze technique was later extended to the measurement of reading comprehension in the native language. The score of individual test subjects is then taken as an estimate of their reading comprehension ability. Cloze tests have also occasionally been presented in auditory form in order to measure listening comprehension in the native language (e.g. Peisach 1965). More recently, a number of studies have dealt with the use of the cloze technique in measuring foreign language proficiency. A survey of cloze test research is given in Oller (1972).

As reported by Oller (1972), the cloze technique appears to be extremely useful in measuring foreign language proficiency. Studies of cloze tests have yielded impressive reliabilities and item analysis results (Oller, op.cit., p. 13). High correlations have been found between scores in (written) cloze tests and other measures of foreign language proficiency (Oller, op.cit., p. 14). It is interesting to note that the highest correlations have been found with tests of listening comprehension, including dictation. This may seem surprising, since the mode of presentation is completely different in the two testing situations and appears to make different demands on test subjects:

> For one thing, phonology is presumably an important factor in listening comprehension though it should not be a significant factor in taking a cloze test. For another, it is frequently assumed that written tests allow unlimited time for processing whereas orally presented materials invoke short-term memory constraints. (Oller, op.cit., p. 20)

Oller attempts to explain the remarkable correspondence between the two types of tests by referring to a "grammar of expectancy" as "the chief mechanism underlying the skills of thinking, understanding, speaking, reading and writing" (p. 18). Listening and reading — and filling in blanks in cloze tests — are processes which require a cooperation between the productive and receptive skills of language:

> The information provided in the cloze test allows the student by analysis to synthesize a greater whole ... By sampling the information that is present the subject formulates hypotheses or expectations, about information that is to follow. By sampling subsequent sequences, he either confirms or disconfirms these expectations. If the expectations are disconfirmed they must be revised and new hypotheses must be formed. (Oller, op.cit., p. 19)

The hypotheses are formed on the basis of the available information and the test subject's knowledge of the language. Compare the account of speech perception in Fry (1970):

> When we take in a spoken message, it is necessary for our ears to receive sound-waves originating with the speaker ... These sound-waves and the information that the brain is able to extract from them form no more than a rough guide to the sense of the message, a kind of scaffolding upon

> which the listener constructs or reconstructs the sentences originating in the speaker's brain. (p. 31)
>
> The decoding of a spoken message is then a twofold process in which the listener receives acoustic information and combines it with linguistic (including statistical) information. (p. 32)

It is this high-level correspondence between cloze tests and tests of listening comprehension which accounts for the high correlations found by several investigators.

Cloze tests and tests of listening comprehension, including dictation, appear to measure basically the same ability in test subjects. There is, however, an important limitation of the cloze technique. Carroll (1972, p.18) points out that cloze scores are chiefly dependent on the "local redundancy" of a passage, i.e. "the extent to which linguistic cues in the immediate environment (generally, in the same sentence) of a missing word tend to supply it". He reports further that it is "even possible to secure cloze scores on the basis of meaningless material so long as grammatical cues are present" (p. 19). A possible way of overcoming the "local redundancy" problem would be to introduce multiple word deletions. In this case the restoration of the missing elements would have to rely much more on the comprehension of the whole text. Oller (1972, p. 6) mentions the possibility of using multiple word deletions but states that this variation of the cloze technique has not been investigated.

Let us now examine partial dictation against the background of the above account of the cloze technique. Partial dictations are similar to cloze tests in the sense that students are provided with

a written text with blanks to be filled in.[3] There are, however, some important differences:

a) Multiple word deletions are normal in partial dictation, as defined in this paper.

b) The students listen to the whole text, including the missing portions.

By using multiple word deletions, the "local redundancy" problem referred to above is solved. By letting the students listen to the text we can be sure that we measure comprehension proper, not inferential processes.[4] It might be objected (1) that the test becomes too easy, and (2) that we test a purely passive skill in this way. However, it is possible to increase the level of difficulty of the test by choosing recordings in an unfamiliar dialect or containing difficult words and constructions or recorded with a high speed of delivery. Moreover, we hope to have demonstrated already that aural perception is not a passive skill but involves the active use of all aspects of the language system.

3 PARTIAL DICTATION

Partial dictation is certainly not a new technique. It has been recommended as a useful exercise (e.g. Black 1969,342) and has probably been used by many foreign language teachers. Valette (1967) mentions partial dictation as a language proficiency test, although she restricts the portions to be filled in by the student to "function words, or often merely prefixes or endings" (p. 139). There does not appear to be any detailed account of partial dictation, as defined in this paper, in the literature. This paper is based on a close investigation of a partial dictation administered recently at the Department of English, Lund University, Sweden. A description of the test, which consisted of two separate passages of English text, appears in Appendix 1.

3.1 RELIABILITY

The reliability of partial dictations appears to be very high. A product-moment correlation of .88 was found between the scores in the first and second passages of the test described in this paper. The figure is based on the scores of all (80) participants in the test. The correlation is remarkably high considering the fact that the two passages (D1 and D2) differ in at least three respects:

a) D1 was constructed according to procedure II, D2 according to procedure I (cf. Section 3.3).
b) D1 was recorded by a native speaker of American English (Inland Northern). The speaker's pronunciation in D2 was typically British (Received Pronunciation). The American speaker was female, the British speaker male.
c) D1 was undistorted but was recorded with a

very high speed of delivery. The speed of delivery was lower in D2; on the other hand, there were slight disturbances on the tape. D2 was taken from a British radio programme. The high correlation between scores in D1 and D2 further illustrates the fact that comprehension is largely independent of the mode of presentation of a test (cf. Section 2.2). A practical conclusion that could be drawn from this result is that exercises in listening comprehension should concentrate more on higher levels of language structure than on phonology.

3.2 VALIDITY

In an early experiment (see Johansson 1973) high correlations were found between a partial dictation and two other types of tests: a test in grammar and vocabulary (.70) and a listening comprehension test consisting of questions on a recorded text (.83). The scores of 20 randomly selected students taking the test described in this paper were compared with scores in the following tests:

a) A test of English vocabulary consisting of 120 multiple-choice items. The student has to choose the correct Swedish translation from among five alternatives presented. The English words are chosen on the basis of a word frequency list.

b) A test of English grammar consisting of 100 multiple-choice items. Each item consists of an isolated English sentence containing a grammatical choice situation. The student has to choose the correct alternative from among two or three presented.

c) A test of English pronunciation conducted in a language laboratory. The student has to read (1) a series of words and sentences

testing different phonemic contrasts and (2) a series of words and phrases designed to test the pronunciation of individual lexical items. Test items are chosen on the basis of a word frequency list.

d) A test of Swedish vocabulary consisting of 100 multiple-choice items. The student has to choose the closest Swedish synonym or definition from among five alternatives presented. The test items are selected on the basis of a word frequency list.

e) A speaking test conducted in the language laboratory. The student has to re-tell a story and describe a series of pictures. The number of serious mistakes in grammar and vocabulary constitutes the score in the test.[6] (The correlation in this case is calculated on the basis of the scores of the total number of students, 36, taking part in the test.)

The following product-moment correlations were computed:

English vocabulary	English vocabulary + English grammar	English pronunciation	Swedish vocabulary	Speaking test (English)
.80	.86	.78	.65	.62

The results should be compared with the figures reported by Oller (1971). The rather high correlation between partial dictation and the test in Swedish vocabulary may seem surprising. On the other hand, the correlation between the latter test and the English grammar test was still higher (.70). The correspondence indicates that there is a relationship between native and foreign language proficiency.[7]

Valette (1964) concludes that dictation is a valid test of overall language skills only for students with little practice in dictation and that emphasis on such practice develops proficiency in that skill alone without leading to proportional proficiency in other aspects of language learning. We have not been able to establish whether this is also true of partial dictation, since all the students taking part in our tests have had comparatively little practice. This possible limitation of the testing technique should, however, be kept in mind.

3.3 CONSTRUCTION AND ADMINISTRATION

The procedure for constructing partial dictations is fairly simple. There are two possible ways of doing it.

I a) Use a recording as a starting-point, e.g. from a radio or television programme.
b) Make a written version of the recording.
c) Decide what portions should be left out in the student's answer sheet.
d) Insert pauses in the recording for each missing portion in the text.

II a) Select a suitable text to be recorded.
b) Have the text recorded by a native speaker of the language.[8]
c) Decide what portions should be left out in the student's answer sheet.
d) Insert pauses in the recording for each missing portion in the text.

Alternative II was used in the first passage of the test analyzed in this paper, alternative I in the second (see Appendix 1).[9]

We shall not deal with all the problems involved in finding the suitable level of difficulty of the text/recording used as starting-point. This will obviously have to vary depending on the level of proficiency of the group of students which the test is designed for. The test described in this paper was designed for Swedish university students of English, and the degree of difficulty is consequently high.

The most important step in constructing the test is deciding what portions should be left out in the written version of the recording appearing in the student's answer sheet. Earlier experiments have shown that, if possible, portions should be left out immediately preceding a possible pause.[10] Usually, students have been asked to fill in missing words at the end of clauses or sentences. The length of the missing portions may vary depending on the possibilities offered by the text. It remains to be established if there is an optimal length.[11] In the test described in this paper the number of words left out in each case varies between 10 and 2. An examination of the results for the different test items shows that they only very rarely failed to discriminate between the top and the lower half of a group of 20 randomly selected students (see Appendix 2). In no case did the top half receive a lower average score than the lower half.

Inserting pauses in the recordings involves no particular problem. Since the length of the missing portions varies, no standard pause length can be used. It deserves to be pointed out that the pause must follow immediately after the missing portion in the text. Test items of the following kind have been found to be unsatisfactory (the words to be filled in by the student are underlined; the location of the pause in the recording is indicated by //):

Britain has the fastest rising cost of living of any other country in the world.//

At this moment, Richard Nixon stands within an ace of victory in the 1968 presidential election.//

In cases like these students appeared to have great difficulty in remembering which words they had to fill in. Omissions were also especially frequent in test items of this kind. If there has to be a missing portion in the middle of a sentence, it must therefore be immediately followed by a pause, e.g.:

Britain has the fastest rising cost of living // of any other country in the world.

At this moment, Richard Nixon stands within an ace of victory // in the 1968 presidential election.

Even though there is a break in the middle of the sentence, test items of the latter kind have worked satisfactorily.[12]

The administration of the test should preferably be carried out in a language laboratory in order to ensure equal listening conditions for all participants.[13] If many groups of students have to take the same test, it is advisable to construct several passages for partial dictation and select a different combination of passages for each successive group. This is possible because of the high correlation between partial dictations based on different texts (cf. Section 3.1).

3.4 SCORING SYSTEM

The scoring system used so far in our partial dictations

has been very simple: 1 point per correct word.[14] In order to minimize the importance of spelling, certain spelling errors have been disregarded. The guiding principle has been to disregard errors which would not affect pronunciation, provided that the word is clearly recognizable and distinct from other words with a similar spelling. Examples:

strycnine	for	strychnine
poisened	for	poisoned
repeate	for	repeat
naturralist	for	naturalist
speciman	for	specimen
passionatly	for	passionately
straggeling	for	straggling
exibits	for	exhibits
telescopique	for	telescopic
rifels	for	rifles
amunition	for	ammunition
blizards	for	blizzards
convays	for	conveys
elequence	for	eloquence
litterary	for	literary
reasonibly	for	reasonably
sentimentel	for	sentimental
frosen	for	frozen

On the other hand, words containing errors of the following kind have been marked as incorrect:

strickney	for	strychnine
speciment	for	specimen
repit	for	repeat
forcans	for	falcons
causal	for	casual
passinetly	for	passionately
exapit	for	exhibit
ecually	for	equally

hear	for	here
rifals	for	rifles
aminition	for	ammunition
blizers	for	blizzards
exitment	for	excitement
allequence	for	eloquence
tempitures	for	temperatures
capiture	for	capture
inaxerable	for	inexorable
where	for	were

In doubtful cases, there has been a tendency to reject frequent words containing spelling errors, whereas the range of accepted spellings has been slightly greater for more unusual words. Words containing the wrong derivational or inflexional ending have always been marked as incorrect.

There is no doubt that the scoring system could be improved considerably, if the degree of acceptability of student responses is taken into account. Consider, for example, errors of the following kind:

inexorable	inexable
	in extra
	an extra board
	inevitable
ledger-space	legislates
	leggaspace
	leisure space
	living-space

It would be desirable to reward syntactic and/or semantic compatibility of student responses (e.g. inevitable and living-space) as against omissions or responses which merely seem to reflect a vague phonological impression; cf. the discussion in Section 4.2 as regards incorrect responses of students with a higher versus lower score in the

test. Grading the degree of acceptability of responses would, however, make the scoring far more difficult and probably also introduce a greater amount of subjectivity into the test. In view of the satisfactory results obtained with the present scoring system (cf. above), a change appears to be unnecessary.[15]

4 ERROR ANALYSIS

Error analysis is currently a very active field of applied linguistics. However, the analysis is generally limited to the investigation of errors in written and spoken production. Errors in perception have only occasionally been dealt with,[16] no doubt due to the problems involved in the testing of listening comprehension. We shall give a brief and incomplete account of errors in the test described in this paper in order to (1) further illustrate what is being tested in partial dictation and (2) demonstrate how such an analysis may reveal perception problems among learners with a particular language background (in this case, Swedes learning English) as well as throw light on general aspects of speech perception.[17] The error analysis is based on the responses of 80 students.

4.1 TYPES OF ERRORS IN PARTIAL DICTATION

Errors in partial dictations may be classified in various ways. An obvious basis of classification is to group errors according to type of deviation from the text: omissions, additions, reorderings, substitutions, etc. Another possible alternative is to classify them according to the presumed psychological mechanisms involved: anticipation errors, perseverance errors, etc. The primary classification chosen here is based on the level of linguistic structure which the error may be referred to: phonology, lexicology and grammar.

4.1.1 Phoneme errors

In examining phoneme errors it is important to bear in mind that such errors can only be inferred indirectly from the often uncertain evidence afforded

by spelling. We shall limit the discussion to reasonably clear cases. The following types will be distinguished: errors of omission, substitution, addition and reordering.

4.1.1.1 Omission

Omissions, as well as other types of phonemic errors, seem to be especially frequent in unstressed syllables. A few examples may suffice to illustrate the point:

inexable	for	inexorable
purnent	for	pertinent
rescenly	for	reasonably
telescopy	for	telescopic
passionly	for	passionately
strickney	for	strychnine

One case of omission, which occurred repeatedly in stressed as well as unstressed syllables, should be noted in particular, i.e. the omission of /r/. Examples (types and tokens):

passier, passur etc.(12)	for pressure
passage (10)	for pressure
tempitures (2)	for temperatures
inexable, inaxable etc.(5)	for inexorable
stickny, stiqwin etc.(8)	for strychnine
couw	for crow
ruel, rule etc.(7)	for rural
cool, coolty etc.(12)	for cruelty
micombe, micion	for migrant

The following errors should possibly also be interpreted as involving omission of /r/:

purpose (2)	for propose
perviding (2)	for providing

It should also be mentioned that final /t, d, s, z/

representing the endings -(e)s and -ed were frequently omitted. This may in part be a sign of grammatical weakness and will be dealt with in Section 4.1.3.4.

4.1.1.2 Substitution

Substitutions form the most frequent type of phonemic error. Like omissions they are especially frequent in unstressed syllables. Examples:

adverose	for	admiral's
legislates	for	ledger-space
westwan	for	westward
inextrabur	for	inexorable
erving	for	urban
purtirunt	for	pertinent

In order to limit the material we shall mainly deal with substitutions in stressed syllables. We shall only consider types of substitution which appear in several test words or appear frequently in one particular test word. For a summary of consonant and vowel substitutions, see Tables 1 and 2. (pp. 22-23)

Apart from the substitutions summarized in the tables, the following cases are of special interest:

a) the substitution of a back rounded vowel for dark /l/, as evidenced by:

an extra bow, un extraroad,
an extra board, an extra boarn,
an extra bone, an extrabow for inexorable

The following example seems to show the opposite tendency, i.e. the insertion of /l/ in place of or next to a back rounded vowel:

fuels (5), feels (5),
fuel, feel for fuse

b) the substitution of nasal + homorganic stop for a combination of nasal + homorganic stop with

another place of articulation:

intrigated, justicated, indicnated, indicated	for impregnated
contemporal, contempious	for potential
micombe	for migrant

Notice also the substitution of a nasal homorganic with a following stop in:

secund, secound	for succumbed

In this case the change of articulation results from a failure to assign a correct morphemic analysis.

In the passage read by the American English speaker the following additional substitutions are worth noticing:

a) A final unreleased /p/ was interpreted as a /b/ in:

stumb (7), stamb (2), slumb	for stump

An isolated instance of stum was also found.

b) Voiced /t/ was interpreted as a /d/ in:

shoud, shud and various other responses (14) containing /d/	for shot

Most of the substitutions dealt with appear to be independent of neighbouring sounds. They might be termed paradigmatic substitutions. Clear examples of syntagmatic substitutions are infrequent:

chapture, chapter	for capture
distant short (1), distance shores (7)	for distant shores

Errors of this kind are related to the type of error to be discussed in Section 4.1.1.4.

Table 1. Consonant substitutions

Consonant contrast	Test word	Student responses (types and tokens)
/v/ - /w/	vast	waist, western, waste, worst (2), wast
	conveys	conwaised, conweist
	eloquence	elekvents
/r/ - /w/	westward	westrode
	rifles	wifels
	run out of	awan with
/s/ - /z/[18]	conveys	convest (2), convinced (2), conveiced, conweist, conviced (5), convessed (2), convaiced, conveyced, convace, convainced, conveist, convaist, convaced, convast
	blizzards	blassets, blisters, blessives
	reasonably	recently, rescently
	his	this (3)
	shores	source, sauce, chource
	sends	sents
/s/ - /tʃ/	pierced	pearched (7), pitched
/s/ - /ʃ/	pressure	passage (10), passer etc. (5)
	shores	sauce, source
/ʃ/ - /tʃ/	naturalist	nationalist (5)
	natural	national (3)
	shores	chource
/ʒ/ - /dʒ/	ledger-space	leisure space (15), leasure space (2)

Table 2. Vowel substitutions

Vowel contrast	Test word	Student responses (types and tokens)
/i/ - /e, æ/[19]	stiff	dead (3), death, stef, step, skaf, steff
	blizzards	blesers, pleasures, blassets, deserts, blessives, pleasant, efforts
	strychnine	streptine, stateny, stecking
	skill	swell
/e/ - /æ/	eloquence	allequence etc. (5)
	pressure	passier etc. (9), passage (10)
	inexorable	inaxable etc. (5)
	ledger-space	ladder-space
	than	then (2)
/ʌ/ - /ɔ/	lusts	loss (3), losts (3), lost (15)
/e/ - /ei/	bait	bet (3), bed, death, leg, beg

4.1.1.3 Addition

The addition of phonemes normally occurs together with substitutions and may be hard to distinguish from pure substitutions. Examples:

dequal	for	decoy
blisters	for	blizzards
fuels	for	fuse
pingeon	for	pigeon

We shall deal with one type of addition which was found repeatedly in the test examined, i.e. the incorrect addition of /t/ after /n/. It occurred in the following cases:

speciment (11), spacement (2) spicement, stasch ment, speciement	for specimen
eliquants, elequantes, eliquent, elequants (2), eloquants, elekvents, eloquents (2), eloquent, allequents, alacwents, elephants	for eloquence
pigent	for pigeon

Examples like these may indicate that combinations of nasal + homorganic stop function as units; cf. the substitution of combinations of nasal + homorganic stop referred to in the preceding section.

Additions are frequently caused by syntagmatic circumstances, e.g. the addition of /n/ in pingeon for pigeon. This is especially true of the incorrect addition of /t, d, s, z/ representing the inflexional endings -ed and -(e)s; see Section 4.1.3.4 below.

4.1.1.4 Reordering

Sometimes the order of phonemes is inverted. This type of error is related to the cases of syntagmatic

substitution and addition referred to in the preceding sections. Examples:

tradegi	for	tragedy
aminution	for	ammunition
distans short	for	distant shores

More complicated instances are:

purtants infacts	for	pertinent facts
stigeron peregrine	for	straggling pigeon
dust and lastes	for	less than justice

The following cases involve omission as well as reordering:

havened succome	for	having succumbed
ensolver	for	involved observer

We shall not deal further with reorderings, since they appear to be comparatively rare.

4.1.2 Lexical errors

Phoneme errors normally result in lexical errors, i.e. a failure to specify lexical items. Phoneme errors appear, however, frequently to be caused by lexical weakness. It is certainly no coincidence that the substitution of /e, æ/ for /i/ occurred particularly in the comparatively rare words blizzards and strychnine,[20] whereas no case of such substitution was found in e.g. distant and literary. This means that the identification of phonemes is partly based on higher levels of language than phonology.[21] In this section we shall briefly survey the following types of lexical error: omission, substitution and addition.

4.1.2.1 Omission

Omission of lexical items forms the largest category of errors. As is to be expected, omissions occur

especially in the case of comparatively rare words, e.g. blizzards, slumped, inexorable. Omissions appear further particularly in the middle of a phrase or sentence. The words most frequently omitted in rather more closely related to the literary taste of 1972 were: closely, related and literary. Another factor which seems to be involved in omission is stress. Everything else being equal, a content word is more likely to be omitted, replaced or distorted the weaker its stress is. A further factor which might be significant is the grammatical function of words. It is conceivable that modifiers would be more likely to be omitted than heads of constructions. The evidence is not clear on this point; cf. e.g. the examples quoted in Section 4.1.3.3, the end.

4.1.2.2 Substitution

As in the case of omissions, substitutions are more frequent if the lexical item is unstressed and/or appears in the middle of a phrase or sentence. This may be illustrated by errors in rendering the date 1972 (types and tokens):

> 1922 (2), 1962 (2), 1932, 1902, 1872,
> of the 19th century too

As is apparent from the introductory discussion (Section 4.1.2), substitutions are especially common in the case of infrequent words.

In studying substitutions we may find a scale of phonological and/or semantic approximation to the test word. Examples (the test words are listed to the left):

migrant	migron	bike	falcon
	micron	vision	hawk
	micro	sight	creature

crow	craw	crown	bird
	couw	crop	hawk
	pur	crab	falcon

Errors in the first column indicate an attempt to give a purely phonological rendering of the test word. Responses in the second column are actual English words but semantically inappropriate; those in the third column are English words semantically related to the test word. As regards the connection between these types of responses and the level of proficiency of the student, see Section 4.2.

The following observations may be made as regards the substitution of a lexical item for another:

a) An infrequent word is usually replaced by a more frequent word. Examples:

stumbled, stepped, stopped	for slumped
strangling, struggling, striding	for straggling

b) A word is sometimes replaced by another word that is related to the total context. Example:

hawk, falcon	for crow and migrant

Errors of the latter kind, as well as errors involving the substitution of words which are semantically appropriate but phonologically incorrect, demonstrate that the comprehension of the total context is involved in partial dictation.

In certain substitutions a lexical item is segmented incorrectly (e.g. of the server for observer). Errors of this kind will be discussed in Section 4.1.3.1 below.

In some cases substitutions are clearly syntagmatic. Examples:

just injustice	for	less than justice
observed observer	for	involved observer
skillity	for	skill (in skill and clarity)
excitement and distinctment	for	excitement and distinction

These instances are related to the reorderings dealt with in Section 4.1.3.2 below.

4.1.2.3 Addition

The addition of content words, apart from the cases of incorrect segmentation to be dealt with in Section 4.1.3.1, is infrequent. Examples:

the less passionetly less involved	for	the less passionately involved
for the peregrine falcons	for	for the falcons

The addition of function words is the most common subcategory; this will be dealt with in Section 4.1.3.5 under grammatical errors.

4.1.3 Grammatical errors

The following categories of grammatical errors will be briefly dealt with: errors of segmentation, reordering, restructuring and errors in morphology and the use of function words.

4.1.3.1 Segmentation

Errors of segmentation is the category which perhaps most clearly demonstrates the active use of grammar in partial dictation. This justifies a more detailed presentation of such errors than will be provided for the other categories. Examples:

at be lossy temporitys, at blows its temperature, often lost his temper	for	at below zero temperatures
rather ammunition, run avery nision	for	run out of ammunition
rescently alarmed	for	reasonable well-armed
rightful	for	rifle
unintemporary, untemporary, on the temporary (5), on a temporary (2), on it temporary, on detemporary	for	only temporary
west word, west wood, west world etc. (many examples)	for	westward
of appearance	for	of the period
person in fact	for	pertinent facts
lasting injustice, last injustice (2), just injustice	for	less than justice
having to come (4), having so fun, having so come	for	having succumbed
bright green blue	for	brightly plumed
of the server	for	observer
terse scopit	for	telescopic
cool free beauty, cools, fear and beauty, ccoled beauties	for	cruelty and beauty
over a million	for	urban millions
rule of our caboure	for	rural Arcadia

Examples of incorrect segmentation of inexorable are given in Section 4.1.1.2. As demonstrated by the examples, the correct segmentation is far from given but must be discovered by the student in the continuous chain of sound.

Segmentation errors share certain characteristics with substitutions (cf. Section 4.1.2.2):

a) They occur especially in unstressed syllables.
b) There is a tendency to replace rare words by more frequent ones.
c) As in substitutions, we may observe false segmentations with varying degrees of phonological and syntactic correctness.

Two cases of presumably incorrect segmentation (and substitution) were especially frequent:

a) In the test item conveys the vast genocidal tragedy of the period 44 students wrote a verb form in the past tense (conveyed, convessed, convinced, convised etc.).
b) In the test item lacks the sweep and the eloquence 16 students wrote a verb form in the past tense (usually lacked or liked).

In the first case, the definite article following the verb probably provided the basis for the past tense form of the verb appearing in student responses,[22] though contextual factors may also be involved (the preceding sentences were in the past tense). This has also happened in a minority of the responses in the second case; however, a majority of students (46) have here simply omitted the inflexional ending of the verb (the usual response was lack or like).

4.1.3.2 Reordering

Pure reordering errors are very infrequent. Examples:

his presentation and arrangement of the facts	for his arrangement and presentation of the facts

the dedicated equally sportsman	for	the equally dedicated sportsman
a cross between a brightly plumed pigeon and a crow	for	a brightly plumed cross between a pigeon and a crow

Reorderings occur more frequently together with substitutions and other types of deviations from the text, especially in long test items. Examples:

the less sportsman proposed	for	the less passionately involved sportsman
rather more litterally related to the taste of 1972, rather more literally related to 1972, rather more literary near to the taste of 1972	for	rather more closely related to the literary taste of 1972
to the taste of literary of 1962	for	to the literary taste of 1972
casul rural dream of some rude Arcadia, cassue roaral dream of som golden acadia	for	casual urban dream of some rural Arcadia

In the last-mentioned test item urban in urban dream was repeatedly replaced by rural. However, only one instance was found of the replacement of rural in rural Arcadia by urban.

4 1 3.3 Restructuring

In errors of restructuring the grammatical construction is changed. Restructurings normally occur together with substitutions and/or additions. Examples:

had belowed several temperatures	for	at below zero temperatures

his arrangement in presenting the facts	for	his arrangement and presentation of the facts
I am not proposed to	for	I don't propose to
that distance calls	for	from distant shores
that provides	for	of providing
the less passionately of all insolved	for	the less passionately involved observer
the sadest kind of tragedy	for	the vast genocidal tragedy
frozen stiff as they stumbled over a stump	for	frozen stiff and slumped over a stump

The restructuring in the last example occurred in the papers of many students (as he/they slumped/stumbled etc. over a stump/stumb etc.).

Errors of restructuring often involve changes in segmentation; examples have already been given in Section 4.1.3.1. Restructurings are further frequently combined with reorderings; for examples see the preceding section. Another category of restructuring may appear when test items, especially long test items, are "abbreviated". Examples of such abbreviations:

for the urbans (3)	for	for the urban millions
the less passioned observer, the less impationate observer, the less passionated observer etc. (many examples)	for	the less passionately involved observer
close to (many examples)	for	closely related to
succumbed to his poisonbate/to his poison	for	succumbed to his poisoned bait

4.1.3.4 Morphology

Under this heading we shall deal with errors which could be corrected by changing the grammatical form of a word. Errors are especially common in the use of the inflexional ending -(e)s. Examples of omission or addition of -(e)s:

he send, a falcon swoop, he exhibit, he lack/like	for he sends, a falcon swoops, he exhibits, he lacks
admiral son	for admiral's son
cruelty and beauty feels/flows/goes/fuses/ fuels/rules etc.	for cruelty and beauty fuse together

In the last test item quoted the final consonant of the stem fuse was very often interpreted as a combination of another verb + the ending -(e)s. The following cases of addition of -(e)s, which were very frequent, should no doubt be interpreted as anticipation errors (types and tokens):

his arrangements and presentation of the facts (26), presentations of the facts (4)	for his arrangement and presentation of the facts
as potential baits/ bates for the falcons (13)	for as a potential bait for the falcons

Another case of addition deserves to be noted, since it occurred very often (42 examples):

eye for details	for eye for detail

The probable explanation for this addition is that the corresponding Swedish expression would contain a noun in the plural.

Apart from omissions and additions of -(e)s, the following cases of incorrect use of endings were noted:

a) The form of adjectives and adverbs was repeatedly rendered incorrectly. Examples (types and tokens):

reasonable well armed (14)	for reasonably well armed
passionate involved (13)	for passionately involved
equal dedicated (5)	for equally dedicated
bright plumed (4)	for brightly plumed
was temporarily (2)	for was temporary
sentimentaly or sensation	for sentimental or sensational
naturally history	for natural history
an unusually migrant	for an unusual migrant

b) A case of incorrect use of -ed was dealt with above under segmentation (Section 4.1.3.1), i.e. the use of conveyed etc. for conveys and lacked etc. for lacks. The addition of -ed also occurred frequently in the following instance:

fused/are fused/is fused/refused/forced together (11)	for fuse together

As in the case of conveyed and lacked etc., the addition is probably syntagmatically conditioned. The omission of -ed occurred repeatedly only in having succumbed to his poisoned bait, probably conditioned by a combination of phonological (the following to) and lexical factors (failure to identify succumb).

Apart from cases of incorrect use of endings, there were also, though less frequently, errors of the following kind (types and tokens):

did not ran (6), did not had	for did not run
shoot (4)	for shot
would be compared (16)	for will be compared

The last-mentioned instance seems to be an example of an anticipation error.

4.1.3.5 Function words

Function words are normally unstressed and frequently strongly reduced. Hence, we can expect to find many errors in the use of such words. The principal categories are: omission, addition and substitution. Examples:

a) omissions:

having succumbed his poisoned bait	for	having succumbed to his poisoned bait
as potential bait	for	as a potential bait
I don't propose repeat/ repeating	for	I don't propose to repeat

b) additions:

his arrangement and his presentation	for	his arrangement and presentation
from the distant shores	for	from distant shores
the literary taste of the 1972 (several examples)	for	the literary taste of 1972

c) substitutions:

the elderly admiral's	for	an elderly admiral's
the excitement	for	their excitement
at/to/in continuous blizzards	for	through continuous blizzards
for/to/with/as for potential bait	for	as a potential bait
having succumbed of/ with	for	having succumbed to
eye of/on/with detail	for	eye for detail

Substitution errors are especially frequent. Partial dictation would seem to be a particularly efficient means of testing the command of function words.

4.2 LEVELS OF COMPREHENSION

We have seen earlier (cf. especially Sections 4.1.2.2 and 4.1.3.1) that the degree of phonological correctness and syntactic/semantic acceptability may vary in incorrect responses. In this context we might talk of different levels of comprehension:

I The lowest level is represented by responses which seem to reflect an analysis on the phonological level only. Example:
legislades for ledger-space

II On a higher level, the chain of sound has been analyzed morphologically and/or lexically, though the response is syntactically and semantically incorrect. Example:
legislates for ledger-space

III On the next level, the response conforms to the syntactic rules of the language, but it is semantically inappropriate. Example:
ladder space for ledger-space

IV On the highest level, the response is syntactically correct and semantically more appropriate, though it deviates from the text on the level of phonology. Example:
living-space for ledger-space

These levels of comprehension correspond approximately to the stages of error correction outlined in Fry (1964); cf. also the quotation from Fry (1970) given in Section 2.1.

An attempt was made to study the relationship between level of comprehension, as defined above, and the degree of proficiency of students, as measured by their scores in the test. The following three test words, which were very often incorrectly rendered, were selected for study:

pressure	in	westward pressure was inexorable
straggling	in	swoops on a straggling pigeon
ledger-space	in	providing ledger-space for the urban millions

A typical response for each test word and level is listed below together with the number of students who provided a response on that level. The average score of the students is given within parentheses (maximum: 200).

	pressure		straggling		ledger-space	
I	passier	8 (122)	stragelin	5 (126)	legislades	3 (154)
II	pursue	1 (126)	-		legislates	2 (161)
III	Russia	5 (133)	strangling	6 (132)	ladder space	16 (141)
IV	passage	10 (142)	ctruggling	18 (160)	living-space	16 (170)

According to the evidence, there seems to be a clear difference in proficiency between students who have provided responses on lower levels of comprehension and students whose responses belong to the highest level. The evidence is too scanty to allow any conclusions to be drawn as to the mutual relationships between the lower levels. The result of our comparison agrees with the observation made by Angelis (1972) in a report about errors in a transcription exercise that beginning students tended to transcribe with reference to the phonetic level only, whereas more advanced students attempted to produce a transcription which gave a reasonable meaning.

4.3 THE LINGUISTIC EVIDENCE OF COMPREHENSION ERRORS

The aim of error analysis is primarily to reveal the problems of foreign language learners. However, the analysis may not only serve a pedagogical purpose but may be of great theoretical interest. The errors revealed are then used to support or refute various claims about language functioning. This is the approach represented by e.g. Fry (1964), Cohen (1966), Nooteboom (1969), Fry (1969), Bierwisch (1970) and Fromkin (1971). The work carried out by Fry is of special interest in this context, since it deals with the perception of speech. Fry (1970), p. 49, concluded that "... there is compelling evidence in favour of the theory of multi-level decoding to be found in the correction of errors which goes on in speech". This conclusion was based on an examination of errors among native speakers, but it is also borne out by our study of the errors of foreign learners. A detailed analysis of comprehension errors would also make it possible to draw other conclusions about language functioning. We shall list some possible areas of investigation:

a) Perceptual confusion among phonemes. The study should be based primarily on purely phonological renderings of test words (cf. Section 4.2), e.g.
<u>eurgeon</u>, <u>erving</u>, <u>irvin</u>, for <u>urban</u>
<u>urby</u>

<u>micron</u>, <u>micre</u>, <u>migron</u>, for <u>migrant</u>
<u>micombe</u>
Is it possible to set up a hierarchy of relatedness of perceived phonemes? How is this related to a corresponding articulatory hierarchy?

b) The scope of reordering errors. What is the scope of reordering errors and what does this imply about units of coding?

c) Anticipation v. perseveration errors. In the material we have examined anticipation errors

were more frequent than perseveration errors.[23] Would this be true in an examination of a larger material? If so, what does this say about language functioning?

d) Category of errors and psychological characteristics of learners. Is it possible to relate different categories of errors (e.g. errors on different levels of comprehension or the frequency of omissions v. substitutions) to psychological characteristics of learners? What does this indicate about the nature of linguistic competence?

In order to make it possible to draw more general conclusions from studies of this kind, it would be desirable to analyze errors in the same test committed by native speakers of a language as well as by foreign learners of that language with varying source language backgrounds.

5 WHAT DOES PARTIAL DICTATION MEASURE?

It is traditionally assumed that dictation mainly measures sound-symbol correspondences. Cf. e.g.

> "Dictation is primarily a test of spelling." (Somaratne 1957, p. 48)
>
> "Spelling and a few matters of inflection and punctuation can be tested through dictation ..." (Lado 1961, p. 34)

Oller (1971) demonstrated that far more is involved in ordinary dictation than the mere conversion of sound to spelling (cf. Section 2.1). This is especially true of partial dictation, as described in this paper, where the importance of spelling is reduced by the scoring system (cf. Section 3.4).[24] The error analysis in the preceding section clearly demonstrates that the student must use his knowledge on all levels of the language system in order to produce a correct response. The attempt to convert a sound impression to spelling without an analysis on higher levels of language may result in responses of the following kind:

at be lossy temporitys	for	at below zero temperatures
run avery nision	for	run out of ammunition
rule of our caboure	for	rural Arcadia
hav an secund	for	having succumbed

In order to provide the correct response the student must have assigned a correct syntactic and semantic analysis to the test item, i.e. providing the correct response is impossible without comprehension. This is especially true of a partial dictation in English, due to the strong reduction of unstressed syllables and the complicated relationship between sound and spelling.

In tests of listening comprehension it is generally assumed that the student must demonstrate comprehension by providing verbal or non-verbal responses of the following kind: answering questions based on the message, providing paraphrases, following directions, etc. For a survey of different techniques of testing listening comprehension, see Carroll (1972). It is remarkable that accounts of the testing of listening comprehension normally do not include dictation (or partial dictation); cf. e.g. Lado (1961), p. 204-222, Pimsleur (1966), p. 186-194, Valette (1967), p. 47-78, Carroll (1972) and articles in Davies (1968). This is all the more remarkable, since dictation (or partial dictation) seems to be less vulnerable than other testing techniques to the usual criticism of listening comprehension tests, i.e. the fact that they tend to measure intelligence and memory factors as well as "pure comprehension". In his recent survey referred to above, Carroll, p. 3, states that researchers should investigate "whether it would be possible to decrease the correlation of comprehension ability tests with intelligence tests by eliminating or reducing those elements of comprehension tests that call for inferential processes that go beyond sheer comprehension". He further points out that research is needed "to see to what extent it is possible to reduce their dependence on memory" (loc. cit.). Such research into "pure comprehension" should include a close investigation of dictation techniques.

Partial dictation would thus seem to be a useful test of listening comprehension. We do not suggest, however, that it should replace other testing techniques. Other techniques will have to be used to test the higher levels of comprehension mentioned by Rivers (1971) and Carroll (1972):[25]

> We can detect in the process of perceptual construction three stages ... The first

> stage, sometimes called "sensing", is a stage of rapid impressions, only roughly identified and differentiated and is relatively passive and receptive. (Rivers 1971, p. 126)
>
> The second stage is one of identification through segmentation and grouping. (Ibid., p.127)
>
> ... the third stage, that of rehearsal and recoding of the material, which must take place before what we have perceived enters into long-term storage. (Ibid., p. 128)
>
> If the above analysis is correct, testing of comprehension involves consideration of the two conceptually separable stages of the comprehension process. That is, we would like to find out, in a given case, the extent to which the individual "correctly" apprehends the purely linguistic information that is "committed" to the message, and also the extent to which he "correctly" relates that information to some wider context. (Carroll 1972, p. 14)

Besides being used as a listening comprehension test, partial dictation could have another important function. Due to the high correlations with other measures of foreign language proficiency, it might be used as a global estimate of a subject's proficiency in a foreign language; cf. however, the discussion in Johansson (1973), the end.

NOTES

[1] As pointed out by Oller, op.cit., p. 255, these results refute the conclusions of certain authorities who have argued that dictation is an inferior testing technique. Cf. e.g. Lado (1961), p. 34: "Dictation ... appears to measure very little of language."

[2] Note, however, that ordinary dictations may also be administered in this way; see Valette (1967), p. 140 f.

[3] In comparing partial dictation and the cloze technique we only consider the most frequent type of cloze test, i.e. written tests. Note, however, that partial dictation is also different from aural cloze tests (e.g. the types used by Dickens and Williams, 1964, or Peisach, 1965) in that test subjects listen to the complete text and receive a written version with blanks. It would be interesting to study the relative effectiveness of aural cloze tests and partial dictation.

[4] Cf. Carroll (1972), p. 19: "It [the cloze technique] probably depends to a considerable extent on inferential processes."

[5] The choice was random with the exception that students considerably older than the average (over the age of thirty) were excluded.

[6] When the examiner had indicated that a student's recording was too brief, the score was doubled. This scoring system, which differs from that ordinarily used in the test, was used in order to make it possible to compute correlations.

[7] As regards the relationship between native and foreign language proficiency, cf. Stencahl (1972) and Truus (1972).

[8] It is essential to select voices which are unfamiliar to the students taking part in the test. Findings in a recent investigation by Brodkey (1972) suggest that familiarity with the voice of the speaker is a major variable in comprehension.

[9] I am grateful to FK Stig Olsson, Department of English, Lund, for constructing the second part of the test.

[10] It has also seemed preferable not to make any omissions in the first few sentences of each passage (cf. Appendix 1). A possible improvement is to include two or three practice items at the beginning of each passage.

[11] In a survey of work on memory span Miller (1963), p. 210 f., reports that the average college student has a span of about eight digits. The number of words that can be recalled in a connected text is greater. It is suggested, however, that the limiting factor may be the amount of information stored rather than the total number of items. See further Miller (1956).

[12] This could be explained by the fact that the completion of the sentence is given in the answer sheet.

[13] As shown by e.g. Peterson and Barney (1952), test results are influenced by the location of the listener in the room, even though the room is fairly small (9 rows of listeners) and acoustic conditions are favourable. Note, however, that administering a test in a language laboratory makes it more artificial. Actual foreign language speakers are not often heard through earphones. Spencer (1964), who compared test results for groups of listeners who were tested with either earphones or loudspeakers, found that performance for the earphone group was significantly better. Consequently, we must require a high standard of correctness in tests administered in the language laboratory.

[14] A word has only been considered correct if it appeared in the right position in a test item. Compounds, e.g.

ledger-space, have been regarded as two words. On the other hand, numbers have been counted as one word (e.g. 1972)

[15] Different scoring techniques from the one outlined here are described in Valette (1967), p. 137, and Brodkey (1972), p. 208.

[16] Note, however, the investigations by Kühlwein (1970) and Angelis (1972). Cf. also the work by Fry referred to in Sections 4.2 and 4.3.

[17] In this brief survey no consistent attempt is made to identify the causes of the errors. A great number of them, e.g. phoneme substitutions, are due to source language interference. In other cases, e.g. reordering errors, the reason is usually of a more general nature.

[18] The evidence offered by spelling is especially difficult to interpret in this case. Only clear cases have been included.

[19] These errors are of particular interest, since they illustrate that perception and production problems may not be identical. Swedish students have no difficulties with the /i/ - /e, æ/ contrasts in speaking English, though the examples given (and further examples in other dictations) indicate such a problem in perception.

[20] The substitution was also found repeatedly in a frequent word (stiff). The probable explanation for the errors is that this word occurs in a test item where several words were often omitted or incorrectly rendered.

[21] Cf. Fry (1964), p. 173.

[22] This interpretation is supported by the fact that some students wrote the indefinite article instead of the definite article after the verb form: conveyed (etc.) a instead of conveys the.

[23] Note, for example, the substitutions in distant shores (cf. Section 4.1.1.2, the end); the perseveration error distant short occurred once only, whereas the anticipation error distance shores occurred seven times.

[24] The importance of spelling cannot, however, be eliminated completely. For a discussion of comprehension v. spelling errors in dictations, see Kühlwein (1970).

[25] In testing listening comprehension in a foreign language it seems desirable to concentrate on the lower level of comprehension that can be measured by partial dictation. The higher levels mentioned by Carroll and Rivers are probably the same in the native and the foreign language. The relationship between comprehension in the native and the foreign language should, however, be subjected to a detailed examination.

REFERENCES

Angelis, P J, 1972, "Listening Comprehension and Error Analysis". Paper delivered at the 3rd International Congress of Applied Linguistics, Copenhagen, Denmark, 21-26 August 1972.

Bierwisch, M, 1970, "Fehler-Linguistik", Linguistic Inquiry 1. 397-414.

Black, C, 1969, "Aural Comprehension in the Language Laboratory", Moderna Språk 63. 338-346.

Brodkey, D, 1972, "Dictation as a Measure of Mutual Intelligibility", Language Learning 22. 203-220.

Carroll, J B, 1972, "Defining Language Comprehension: Some Speculations", in J B Carroll and R O Freedle (eds.), Language Comprehension and the Acquisition of Knowledge, Washington, D C: V H Winston & Sons. 1-29.

Cohen, A, 1966, "Errors of Speech and Their Implication for Understanding the Strategy of Language Users", Instituut voor Perceptie Onderzoek Annual Progress Report 1, Eindhoven.

Davies, A (ed), 1968, Language Testing Symposium, London: Oxford University Press.

Dickens, M and Williams, F, 1964, "An Experimental Application of 'Cloze' Procedure and Attitude Measures to Listening Comprehension", Speech Monographs 31. 103-108.

Fromkin, V A, 1971, "The Non-Anomalous Nature of Anomalous Utterances", Language 47. 27-52.

Fry, D B, 1964, "The Correction of Errors in the Reception of Speech", Phonetica 11. 164-174.

———, 1969, "The Linguistic Evidence of Speech Errors", Brno Studies in English 8. 69-74.

———, 1970, "Speech Reception and Perception", in J Lyons (ed.), New Horizons in Linguistics, Harmondsworth: Penguin Books. 29-52.

Johansson, S, 1973, "An Evaluation of the Noise Test, A Technique for Testing Foreign Language Proficiency by Perception under Masking Noise". To appear in The International Review of Applied Linguistics.

Kühlwein, W, 1970, "Intra- und interstrukturale Fehlleistungen auf der phonemisch-graphemischen Ebene", PAKS Arbeitsbericht 5. 39-84.

Lado, R, 1961, Language Testing, London: Longmans.

Miller, G A, 1956, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information", Psychological Review 63. 81-97.

———, 1963, Language and Communication, New York: McGraw-Hill.

Nooteboom, S G, 1969, "The Tongue Slips into Patterns". in A G Sciarone et al. (eds.), Nomen: Leyden Studies in Linguistics and Phonetics. The Hague: Mouton. 114-132.

Oller, J W, Jr, 1971, "Dictation as a Device for Testing Foreign Language Proficiency", English Language Teaching 25. 254-259.

———, 1972, "Cloze Tests of Second Language Proficiency and What They Measure". Paper delivered at the 3rd International Congress of Applied Linguistics, Copenhagen, Denmark, 21-26 August 1972.

Peisach, E C, 1965, "Children's Comprehension of Teacher and Peer Speech", Child Development 36. 467-480.

Peterson, G E and Barney, H L, 1952, "Control Methods Used in a Study of Vowels", Journal of the Acoustic Society of America 24. 175-184.

Pimsleur, P, 1966, "Testing Foreign Language Learning", in A Valdman (ed.), Trends in Language Teaching, New York: McGraw-Hill. 175-214.

Rivers, W M, 1971, "Linguistic and Psychological Factors in Speech Perception and Their Implications for Teaching Materials", in P Pimsleur and T Quinn (eds.), The Psychology of Second Language Learning, Cambridge: Cambridge University Press. 123-134.

Somaratne, W, 1957, Aids and Tests in the Teaching of English, London: Oxford University Press.

Spencer, R E and Seguin, E L, 1964, "The Relative Effectiveness of Earphones and Loudspeakers as a Means of Presenting a Listening Test in a Foreign Language", Modern Language Journal 48. 346-349.

Spolsky, B, et al., 1968, "Preliminary Studies in the Development of Techniques for Testing Overall Second Language Proficiency", Language Learning, Special Issue Number 3. 79-101.

Stendahl, C, 1972, "The Relative Proficiency in Their Native Language and in English Shown by Swedish Students of English at University Level". Projektet språkfärdighet i engelska (SPRENG), Rapport 6, Engelska institutionen, Göteborgs Universitet.

Taylor, W L, 1953, "Cloze Procedure: A New Tool for Measuring Readability", Journalism Quarterly 30. 414-438.

———, 1956, "Recent Developments in the Use of 'Cloze Procedure'", Journalism Quarterly 33. 42-48.

Truus, S, 1972, "Sentence Construction in English and Swedish in the Writings of Swedish Students of English at University Level: A Pilot Study", Projektet språkfärdighet i engelska (SPRENG), Rapport 7, Engelska institutionen, Göteborgs Universitet.

Valette, R, 1964, "The Use of the Dictée in the French Language Classroom", Modern Language Journal 48. 431-434.

———, 1967, Modern Language Testing, New York: Harcourt, Brace & World.

APPENDIX 1: THE TEST

The test consists of two passages presented for partial dictation. The first passage was recorded by a native speaker of American English, the second was a recording from the British radio. The test was administered in a language laboratory. Students received the following instructions: "Listen carefully and fill in the blanks in the text. The tape will stop at each //."

The text of the two passages will be given below. Portions to be filled in by students are underlined. For ease of reference, test items are numbered.

D[ictation] 1. Book review: "The Fetterman Massacre" by Dee Brown.

The Fetterman Massacre is the story of Red Cloud's completely victorious ambush of 80 soldiers of the United States army in December 1866. Red Cloud, commanding armies of Sioux and Cheyenne, decoyed a troop of 80 men from Fort Phil Kearny on the Montana road. No one returned alive. The man who took the news of the massacre to the head of the telegraph near Fort Laramie was Portugee Phillips who, (1) at below-zero temperatures and through continuous blizzards//, rode 236 miles in four days. Portugee Phillips survived. The horse died. Captain Brown and Colonel Fetterman, obviously by agreement, (2) shot each other to avoid capture//. The Indians (3) had very few rifles//, the United States soldiers (4) were reasonably well armed//. They did not (5) run out of ammunition//. One officer was found (6) frozen stiff and slumped over a stump//. His body (7) was pierced by 56 arrows//. Red Cloud's (8) victory was complete//. It was, of course, (9) only temporary//. The Montana road was opened again within ten years. (10) Westward pressure was inexorable//. Dee Brown, as in Bury My Heart at Wounded Knee, (11) conveys the vast genocidal tragedy of the period//. He is not

(12) sentimental or sensational//. He draws together (13) the pertinent facts with skill and clarity//. In a sense he lets the facts tell the story, but to say this is (14) doing him less than justice//. It is (15) his arrangement and presentation of the facts// that gives both books (16) their excitement and distinction//. It is possible that some readers may think that Dee Brown (17) lacks the sweep and the eloquence// of the great American historian Francis Parkman but he has an eloquence of his own (18) rather more closely related to the literary taste of 1972//, and I suspect that in the future it will be with writers like Parkman that Dee Brown (19) will be compared//.

D[ictation] 2. Book review: "The Scandaroon" by Henry Williamson.

This book is about some very unusual relationships, not between men and women nor even between men and men, but between men and birds. The workpeople in the small town of Thirby in North Devon keep racing pigeons. That sport has made them deadly enemies of the peregrine falcons, which have lived and hunted along that stretch of the Devon coast for a thousand summers. The falcons swoop on the pigeons as they return home but Sam Baggott, keeper of the Black Horse Inn and a dedicated pigeon racer, has his own deterrent against "they bloody hawks". He (1) sends up decoy birds impregnated with strychnine//, and dances for joy when the falcons tumble out of the sky, (2) having succumbed to his poisoned bait//. The Scandaroon itself is a (3) rather unusual migrant from distant shores//, something of a (4) brightly plumed cross between a pigeon and a crow//. Its arrival in Thirby is greated with considerable interest. Sam Baggott (5) lusts after it as a potential bait for the falcons//. The local doctor regards it with great interest (6) as

a natural history specimen//. But the Scandaroon eventually becomes the property of (7) an elderly admiral's young son//. That, of course, is only the beginning of the story, and since it is rather a good one (8) I don't propose to repeat all of it here//. In this short novel, set just after the First World War, Henry Williamson presents us with several contrasting views of the English countryside — that of (9) the dedicated naturalist//, that of (10) the equally dedicated sportsman//, and that of (11) the less passionately involved observer//. Henry Williamson is not a sentimental nature writer. He knows the countryside and its people too well to lapse into the (12) casual urban dream of some rural Arcadia//. He recognizes that (13) cruelty and beauty fuse together// when a peregrine falcon (14) swoops on a straggling pigeon//. In his novels (15) he exhibits a telescopic eye for detail//, and one feels that he cares more too. Our concern for the environment is essentially a social concern. It derives from the social problem (16) of providing ledger-space for the urban millions//.

The students listened to the recordings only once and were given a few minutes for final revision.

APPENDIX 2: DISCRIMINATION OF TEST ITEMS

For each test item the average score is given below for the top and lower half of a group of 20 randomly selected students participating in the test. As regards the numbering of test items, cf. Appendix 1.

D1			D2		
test item	score of top half	score of lower half	test item	score of top half	score of lower half
1	5.2	2.4	1	5.3	4.2
2	5.9	5.1	2	4.1	3.0
3	3.9	3.8	3	5.3	4.2
4	4.0	3.3	4	6.7	5.3
5	4.0	3.1	5	5.7	2.6
6	5.2	3.7	6	4.7	4.1
7	4.7	3.7	7	4.1	2.3
8	3.0	2.8	8	8.6	7.9
9	1.9	1.4	9	2.8	2.7
10	2.7	1.8	10	3.8	3.8
11	5.6	2.8	11	4.5	3.5
12	2.9	2.5	12	6.7	3.4
13	6.4	4.3	13	4.0	3.3
14	5.0	4.3	14	4.4	3.8
15	6.5	5.6	15	6.1	4.7
16	3.8	3.7	16	6.4	4.6
17	5.2	3.9			
18	8.6	4.6			
19	2.9	2.7			